

#112
03-d-d

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001

TIME: 16:03:53

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02232001\I508821A.raw

```

2 <110> APPLICANT: ROULEAU, Guy A.
3 JOOBER, Ridha
4 BENKELFAT, Chawki
6 <120> TITLE OF INVENTION: POLYMORPHIC CAG REPEAT-CONTAINING GENE AND USES THEREOF
8 <130> FILE REFERENCE: 12948-10
10 <140> CURRENT APPLICATION NUMBER: 09/508,821A
11 <141> CURRENT FILING DATE: 2000-05-26
13 <150> PRIOR APPLICATION NUMBER: PCT/CA98/00884
14 <151> PRIOR FILING DATE: 1998-09-18
16 <150> PRIOR APPLICATION NUMBER: 2,216,057 CA
17 <151> PRIOR FILING DATE: 1997-09-19
19 <160> NUMBER OF SEQ ID NOS: 6
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 328
25 <212> TYPE: DNA
26 <213> ORGANISM: Mus musculus
28 <400> SEQUENCE: 1
29 agggcagcca ctttccccag cattcccagt ccttccctac ctctccact tatgccccaa 60
30 cagtgcaggg tgggtggcag ggggcccact cctacaagag ctgcacagca ccatctgccc 120
31 agcctcatga taggcgatg agtgccaatg cgaacctggc tccagggcaa cgggtccaga 180
32 atcttcacgc ttaccagcct ggcgcctlg gctacgagca gcagcagcaa gcacttcaag 240
33 gccgtacca caccaggaa acactccact accagaacct cgcgaagtac caacactatg 300
34 gacagcaagg ccagggtac tgtccacc 328
37 <210> SEQ ID NO: 2
38 <211> LENGTH: 330
39 <212> TYPE: DNA
40 <213> ORGANISM: Homo sapiens
42 <220> FEATURE:
43 <221> NAME/KEY: modified_base
44 <222> LOCATION: (59)
45 <223> OTHER INFORMATION: N = A or G or C or T/U
47 <400> SEQUENCE: 2
W--> 48 tccttcccca cctcctccac ctactcctcc tctgtcccag ggtggtgggc agggggccna 60
49 ctctataag agttgcacag caccgactgc ccagcccat gacaggccgc tgactgccag 120
50 ctccagcctg gccccggggc agcgggtcca gaatttcat gctaccagt cgggcgcct 180
51 cagctatgac cagcagcagc agcagcagca gcagcagcag cagcagcagc aagcccttca 240
52 gagccggcac catgcccagg aaacctcca ttacaaaac ctgcgaagt atcagcacta 300
53 cgggcagcaa ggccagggt actgccagcc 330
56 <210> SEQ ID NO: 3
57 <211> LENGTH: 23
58 <212> TYPE: DNA
59 <213> ORGANISM: Artificial Sequence
61 <220> FEATURE:
62 <223> OTHER INFORMATION: Description of Artificial Sequence:
63 oligonucleotide
65 <400> SEQUENCE: 3

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001

TIME: 16:03:53

Input Set : A:\Pto.amc

Output Set : N:\CRF3\02232001\I508821A.raw

```

66 ggggcagcgg gtccagaate ttc
69 <210> SEQ ID NO: 4
70 <211> LENGTH: 23
71 <212> TYPE: DNA
72 <213> ORGANISM: Artificial Sequence
74 <220> FEATURE:
75 <223> OTHER INFORMATION: Description of Artificial Sequence:
76     oligonucleotide
78 <400> SEQUENCE: 4
79 agcactacgg gcagcaaggg cag
82 <210> SEQ ID NO: 5
83 <211> LENGTH: 6022
84 <212> TYPE: DNA
85 <213> ORGANISM: Homo sapiens
87 <400> SEQUENCE: 5
88 ggaaccagca ggcacaaagg gatgagggag cggaattgc tctgctaaat gcttttgagc 60
89 tgtcaggaaq ggtgaggagt gatgggtggg ggacattggg gaggagctgg caatgggagg 120
90 gggggggggg ggtagctccc cagtgcctgg gcgctgggga gccggttttg cctcccgcat 180
91 cagtggccgt ccttgccaag actcagctgc aggcagatgt ggcagcggaa ttacagagca 240
92 caectccctg acacagaagt tgtcaatatg cgcacagctg gtggggaggg ccaggcgaag 300
93 gggggagcat taagagctgc gggggggagg aggcaggggt gggaggtggg tgggaggggt 360
94 cttcttgagg caaaaggagg tggcccgctt gaatcgctca tctctgccc cctccctgcc 420
95 catctctccc tctctctctt cctctctccc tctctctctt tttcttcca cagataacca 480
96 gcccagatca tgcagctctt tcgagaaaag tglggtttcc atggcaaaac acagaactac 540
97 cagcagacct cgcaggaaac atcacgccta gagaattaca ggcagccgag ccaggccggg 600
98 ctaagctcgg accggcagcg gctgctcgcc aaggactatt ataaccgcca gcttaccggg 660
99 agctatgagg gtggcgctgg caagccctct ggcactgcag ccgcggtggc cgcgcagaaq 720
100 taccaccgag gcagcaaggc cctgcccaca cagcaaggcc tgcaggggag gccggcttcc 780
101 cctggctacg gcytcaggga cagcagcccc taaccaggcc gctatgctgg tgaggagagg 840
102 ctccaggett gggggggccc acagccacca ccccacagc cgcagccact acctgcaggg 900
103 gtggccaaqt atgatgagaa cttgatgaaa aagacagcag tgcacccagc caggcagtat 960
104 gcagagcagg gcgcccaggt gcccttttcg actcactccc tgcagctcca gcagccaccg 1020
105 ccgcccagc agccctctgg ataccccaag ctccaaaggc agaagctgca gaacgacatt 1080
106 gccctccctc tgcctctccc ccagggtacc cacttctctc agcattcccc gtccttcccc 1140
107 acctctctca cctactctct ctctgtccag ggtggtgggc agggggcccc ctctataag 1200
108 agttgcacag caccgactgc ccagccccat gacaggccgc tgactgccag ctccagctg 1260
109 gccccggggc agcgggtcca gaattctcat gcctaccagt cgggcccctt cagctatgac 1320
110 cagcagcagc agcagcagca gcagcagcag cagcagcagc aagcccttca gagccggcac 1380
111 catgcccagg aaacctctca ttaccaaaac ctgcgcaagt atcagcacta cgggcagcaa 1440
112 ggcaggggct actgcccagc ggacgcagcc gtccggaccg cagagcagta ctaccagacc 1500
113 ttccagccca gctccagcca ctacccccgc cgtctcgtgg gccgctcacc ttctacagt 1560
114 tccacaccgt cgcgctgat gccaaacctg gagaacttcc cctacagcca gcagccgctc 1620
115 agcaccgggg ccttccccgc agggatcact gaccacagcc acttcatgcc cctgctcaat 1680
116 cctcccccaa cggatgccac cagctctgtg gacaccagc ctggcaactg caagccctt 1740
117 cagaaggaca agctccttga gaacctgtg tgggatctca gctlgcagag cctcacggcg 1800
118 ctgaccttac aggtggagaa catctccaac accgtccagc agctgctgct ctccaaggct 1860
119 gctgtgccgc agaagaaaag tgtcaagaac ctctgttcca ggacccaga gcagcataaa 1920
120 agccagcact gcagcccgga agggagcggc tactcagccg agcccagag cacaccgctg 1980
121 tcagagccgc cgagcagcac gccacagtc acgcagcgg agccgcagga ggccgactac 2040

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001
 TIME: 16:03:53

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02232001\I508821A.raw

```

122 ctgagcggct cccaggaccc actggagcgc agcttctctt actgcaacca ggcgcgtggc 2100
123 agccctgcca gggtaacacg caactcgaaq gccaaagccc aglccgtgtc cactgtttct 2160
124 gtgacctctc ctgacacat gtccaccaaa tctgacgact ccttccagag cctacacggc 2220
125 agtctgcgcg tgcacagctt ctccaagtgc gtggcgggtg agcgggactg tccgcggctg 2280
126 ctgctcagcg cctgggcaca ggaggacctg gcctccgaga tccctgggct gcagggaagc 2340
127 atcgggtgaga agcccgacaa agcltgggtc gaagcaccga gcctgggtca ggacagcagc 2400
128 aagccacctt tctcgttggg gaaccacagc gcctgcctgg actctgttgg caaagatgcy 2460
129 tggccccggc ctggggagcc ggaggccctg cccgactcct tgcagctgga caaggcgccg 2520
130 aatgccaaag acttcagccc agggctgttt gaagacctt ccgtggcctt cgtacgcct 2580
131 gaccccaaaa agacaactgg tctctctctc ttgtgtacca agccacctt tggggttcct 2640
132 gctccagacc ccaactacgc agctttttgac tgtttcccg acacaaccgc tgcagctca 2700
133 ggggacagcg ccaaccctt tgcctggcca gagaaaaacc tgggggatgc ttgtccagg 2760
134 tggggattgc accctggcga gcttaccag ggctggagc aggggtggaa ggcctcagat 2820
135 ggcctcagca aaggggacac ccatgaggtc tgggctgccc tgggcttcca ggaggaggac 2880
136 cccctggggg tgaagggtgc clcgttgcgc ggggacttca agcaggagga ggtgggtggg 2940
137 gtgaaggagg aggcaggtg gctgtgcag tgcctcgagg tggccaaggg tgaccggtgg 3000
138 ctgagggaca ggcgggactg ctgttccacc gccgacttcg gggacctccc actgctgcca 3060
139 ccacccagca ggaaggagga cctggaagct gaggaggagt actcctccct atgtgagctc 3120
140 ctgggcagcc ccgagcagag gccctggcatg caggaccgcg tgtcacccaa ggcctcactc 3180
141 atctgacca agagaggagt ggaggaggtg ctggactcca aggcgggctg gggctctccg 3240
142 tggccactct caggggagtc cgtcactctg ctgggccta cagtgggcac cagtcctaaq 3300
143 gtcacagact ggtttgagtc ctctctgtca cacatgaagc caggtgaaga ggggctgat 3360
144 ggggagcgag ctccagggga ttccaccacc tgggagcct ctctggccca gaagccaac 3420
145 aagcctgctg tgcctgggga gccctgcga aagaaagagc ctgtgcaagc gggcaaaagc 3480
146 ttacggagcc gtccgggtgca cggggggtg cccgagggcg aggaactccc atgcagggca 3540
147 ccagtgctgc ccaaaagact cttgtccct gaatcctgca caggggcccc ccagggacag 3600
148 atgggaaggc ctggagcccc aggcgggggg gcctcggaag ggtcccccag gatgttact 3660
149 cgttctctca cggccctgag tgagcccgcc acgcccggac ccccaggcct gaccaccacc 3720
150 cctgcacccc cagacaaact ggggggcaag cagcgagccg ccttcaagtc gggcaagcgg 3780
151 gtggggaagc cctaccacca ggcctgcctc agcccaagca acccggcgc cctgcctgtg 3840
152 gctcccgaca gcagcccgat gggctccaaq aaccaaggag cagactcacc cagcacgcct 3900
153 ggcaggagcc agcgtccat gatccttcgg tcacgcacca aaaccagga gatcttccac 3960
154 tccaagcggc ggagccctc tgagggcccg ctccccaact gccgtgccac caagaagctc 4020
155 ctgcacaca gccacttgcc cgcacatcgc aaggtctcca gcagccccc gaaggagggc 4080
156 aggggtgagc agcgggcaag ggtccccaac cctgttgcaq gcagcaagct ctctgacgg 4140
157 cccctccatg cgtcctaaag gaagtgcgct tcatggcgc cggctccccc caagaagcgg 4200
158 aacctggtct tgcggcagc cagcagcagc agcagcaacg ccagtgcaat gggggagatg 4260
159 ggaaggagga gaggcctgag ggttccccc cctcttccaa gaggatgtct tctcccaaga 4320
160 aagccaaagc caccaaaggg aatggcgagc ctgccacaaa gctccacccc ccggagagcc 4380
161 ccattcctgc ccaagctcg cctctcgcca gccttccagg gggccatgaa gaccaaggtg 4440
162 ctgccacccc ggaaggcgcg gggcctgaag ctggaagcca tctgtcagaa gatcacctcg 4500
163 ccagcctca agaagtgc atgtaaagcg ccaggggcct ctctgttaa tctctgagc 4560
164 ccctccttlt ccgacaaaga ccgtgggctc aaggggtgct ggggcagccc agtgggggtg 4620
165 gaagaaggcc tggtaaatgt gggcaccggg cagaagctcc caacttctgg ggtgatccg 4680
166 ttatgcagaa atccaaacca cagatcctta aaaggcaaac tcatgaacag taagaaactg 4740
167 tcttctactg actgtttcaa aaccgaggcc ttacatccc cggaggccct gcagcctggg 4800
168 gggactgccc tggcgctcaa gaaggaggc cggaaaggcc gggcaggggc ccattggactc 4860
169 tccaaaggcc cgttgagaa gggccctat cttggcccg ctctgctcct gactcccca 4920
170 gacagggcca gtggcacaca aggggccagt gaggacaact ctgggtggag aggcagaag 4980

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001
TIME: 16:03:53

Input Set : A:\Pto.amc
Output Set: N:\CRF3\02232001\I508821A.raw

```

171 ccaaagatgg aggaagctggg cccgcccctcc cagcccccgg agggcaggcc ctgccagccc 5040
172 cagacaaggg cacagaaaca gccaggccac accaactaca gcagctatc caagcgggag 5100
173 cgcctcactc qgggcccggc caagaacacc acctcttcac cctgtaaggg gcgtgccaa 5160
174 cgaagacgac aacagcaggt gctgcccctg gatcccgcag agcctgaaat ccgcctcaag 5220
175 tacatttctt cttgcaagcg gctgaggtca gacagccgga ccccgcctt ctcaccctc 5280
176 gtgcgggtgg agaagcgaga cgcgttcacc accatalgca ctgltgtcaa ctcccctgga 5340
177 gatgcgcccc agcccccacag gaagccttcc tcctctgcct cctcttctc atcctcgtcc 5400
178 tcgtttctct tggatgcagc cggggccctc ctggccacac tccttgagg ctccatcctg 5460
179 cagccgcggc cctccttgcc cctctctccc acgatgcact tggggcctgt ggtttccaa 5520
180 gccctgagta cctcttgctt tgtttgctgc ctctgccaaa acccggccaa ctccaaggac 5580
181 cttgggggacc tctgtgggac ctactaccct gaacactgcc tccccaaaaa gaagccaaaa 5640
182 ctcaaggaga aggtgcggcc agaaggcacc tgtgaggagg cctcctgccc gcttgagaga 5700
183 aactcacaag gtcccagtg tgcagctgcc gccactgcc ggaagcccc caggtgacgg 5760
184 ccagctgac ccggccaagc agggcccact gcgcaccagt gcccggggcc tgtcccggag 5820
185 gctgcagagc tgcctactgt gtgctggccg ggaggatggg gccgaggagg cagcccccagc 5880
186 cgacaaggtt cgcacaacatg agtgacgcaa ggaggctccg gcagagcccg gcggggaggc 5940
187 ccaggagcac tgggtgcctg aggcctgtgc cgtgtggacc ggcggcctct acctggtggc 6000
188 cgggaagctc ttgggctgc ag                                     6022
191 <210> SEQ ID NO: 6
192 <211> LENGTH: 22
193 <212> TYPE: DNA
194 <213> ORGANISM: Artificial Sequence
196 <220> FEATURE:
197 <223> OTHER INFORMATION: Description of Artificial Sequence:
198     oligonucleotide
200 <400> SEQUENCE: 6
201 tggccttgc tgcctgtagtgc ct

```

22

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001

TIME: 16:03:54

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02232001\I508821A.raw

L:48 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2